

MAMoC: Multisite Adaptive Offloading Framework for Mobile Cloud Applications

Dawand Sulaiman
School of Computer Science
University of St Andrews
St Andrews, UK
djs21@st-andrews.ac.uk

Adam Barker
School of Computer Science
University of St Andrews
St Andrews, UK
adam.barker@st-andrews.ac.uk

Abstract—This paper presents MAMoC, a framework which brings together a diverse range of infrastructure types including mobile devices, cloudlets, and remote cloud resources under one unified API. MAMoC allows mobile applications to leverage the power of multiple offloading destinations. MAMoC’s intelligent offloading decision engine adapts to the contextual changes in this heterogeneous environment, in order to reduce the overall runtime for both single-site and multi-site offloading scenarios. MAMoC is evaluated through a set of offloading experiments, which evaluate the performance of our offloading decision engine. The results show that offloading computation using our framework can reduce the overall task completion time for both single-site and multi-site offloading scenarios.

I. INTRODUCTION

With the sheer ubiquity of powerful mobile devices, it is now possible to harness the capabilities of these mobile devices in order to execute, and offload compute-intensive applications. Most of the modern mobile devices on the market have quad-core and octa-core CPUs, and mobile devices of today are more computationally powerful than the PCs of the last decade [1].

The concept of using cloud hosted infrastructure as a means to overcome the resource-constraints of mobile devices is known as *Mobile Cloud Computing (MCC)*, and allows applications to run partially on the device, and partially on a remote cloud instance, thereby overcoming any device-specific resource constraints. However, as smart phones and tablets gain more CPU power and longer battery life, the meaning of MCC gradually changes. Instead of being fully dependent on the cloud, a number of nearby devices can be used to coordinate and distribute content and resources in a decentralized manner; this is known as *Mobile Ad hoc Cloud Computing* [2].

Although MMC has helped many application developers to overcome the limited resources of mobile devices, it has also created a new set of challenges, such as the possibility of high network latencies and low bandwidth availability between the mobile device and the cloud. Other research efforts investigated frameworks which allow the cloud to move closer to the user in the form of cloudlets [3] [4]. Cloudlets are trusted, resource-rich computers that are connected to the Internet and available to nearby mobile devices. Other approaches include a group of nearby mobile devices to leverage lower

end devices thus the formation of a local mobile cloud [5] also referred as *Mobile Device Clouds* [6] and *Mobile Edge Clouds* [7]. Mobile devices with less computational power and lower battery life can be empowered by the nearby mobile devices to run resource-intensive applications. Therefore, more efficient and reliable methodologies need to be explored for resource-hungry and real-time applications such as face recognition, data-intensive, and augmented reality mobile applications.

The heterogeneous mobile cloud environment contains different types of computing resources such as remote clouds, cloudlets, and mobile devices in the vicinity that can be utilized to offload mobile tasks. Heterogeneity in mobile devices includes software, hardware, and technology variations. In MCC, providing collaboration among various mobile and cloud nodes with different interfaces is a significant matter. Dynamic environmental changes is one of the challenges facing the offloading decision making in mobile cloud applications. Mobile Cloud frameworks need to adapt to these changes for efficient task partitioning and high QoS of the mobile applications running on end user devices. Existing mobile computation offloading frameworks lack the automated transparency feature so that the surrounding devices can be detected and the computation offloading take place in a seamless manner [8].

This paper presents *Multisite Adaptive Mobile Cloud (MAMoC)*, a unified framework which allows each mobile device within the shared environment to intelligently offload its computation to other external platforms. For the individual mobile devices, it is important to make the offloading decision based on network conditions, load of other machines, and mobile device’s own constraints (e.g., mobility and battery). Moreover, to achieve a global optimal task completion time for tasks from all the mobile devices, it is necessary to devise a task scheduling solution that schedules offloaded tasks in real time. To achieve the vision of mobile computing among heterogeneous converged networks and computing devices, designing resource-efficient environment-aware applications is an essential need. The offloading decision engine needs to adapt to the dynamic changes in both the host device and connected nearby and remote devices. Changes in hardware such as CPU workload, available memory, and battery state and level need to be carefully monitored.

The main contributions of this work are:

- 1) Leveraging mobile applications running on lower end devices with nearby mobile devices, cloudlets, and remote cloud resources using a single framework.
- 2) An API on top of our framework, which allows application developers to use it as a simple programming model to build mobile cloud applications and abstract complex underlying heterogeneous technologies.
- 3) Through a real-world evaluation, we demonstrate that it is possible to speedup computation of mobile devices by using the adaptive offloading algorithm in both single-site and multi-site offloading scenarios.

II. RELATED WORK

Mobile Computation offloading transfers processing from the mobile device to other service providers. Mobile application is partitioned and analyzed so that the most computational expensive operations at code level can be identified and offloaded [9]. The objective is to improve the computation performance, enable advanced functionality, and preserve scarce resources. In the mobile-to-cloud offloading model, most challenges arise from partitioning the mobile application code to remote and local tasks based on the dependencies of each task. As such, the current solutions can be categorized by partitioning technique into static and dynamic. The authors at [10] produce an elastic application model in the form of weblets which can be platform independent or dependent. The decision of offloading is based on contextual components stored in the cloud including device status (CPU load, battery level), performance measures of the application for quality of experience and user preferences. By doing so, the application model supports multiple running modes: power-saving mode, high speed mode, low cost mode or offline mode. Mobile offloading overcomes the resource limitations of lower end devices by splitting resource-intensive tasks and allocating subtasks to other resource-rich devices. Offloading may be performed at different granularities ranging from methods and individual tasks [10] to applications [11] and virtual machines [12].

The mobile cloud framework developed in [13] uses the same interface of existing cloud APIs for the collocated virtual computing providers. This allows a seamless integration with the existing cloud infrastructures. On the other hand, Cuckoo framework [14] uses the native Android partitioning to separate the user interface from the background computational code. This eases the design and implementation of MCC applications, as mobile application developers do not require any cloud computing knowledge, such as integrating with offloading APIs.

Most existing MCC proposals concentrate on single-site offloading [15] i.e., offloading application's parts from the mobile device to a single server. However, as the number of surrounding devices and cloud computing and storage increases, it is more common that an application can be executed on multiple servers [16]. It is shown that we can obtain better performance from multisite offloading. Therefore,

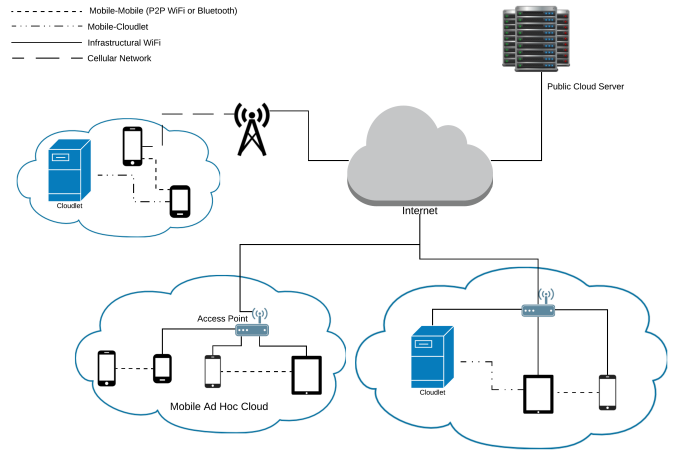


Figure 1. Heterogeneous Multisite MCC Architecture

multisite offloading is considered as a generally realistic model in this work. However, making decision for multisite offloading problem is an NP-complete problem, and hence, obtaining the optimal solution is time consuming. Hence, we use a simple near optimal decision algorithm to find the best-possible partitioning for offloading to multisite clouds/servers.

We have shown in our previous work that offloading tasks from lower end devices to nearby devices can save up to 50% in computation performance [17]. In this work, we propose the creation of a hybrid mobile cloud framework in which the resources are nearby mobile, immobile (cloudlets) and remote cloud servers. MAMoC focuses on reducing the total task computation time and the overall consumed power of the interconnected devices. Similar to [14], we rule in favor of injecting the offloading code transparently to the developer. By doing so, we preserve the mobile application development process, as the developer is unaware of the underlying mechanisms and is only left with the possibility of giving hints about code that may be offloaded. Such a development model encourages the decoupling of components and the modularity of mobile applications.

III. SYSTEM DESIGN

Our objectives for MAMoC include improving the running time of the compute-intensive tasks on the mobile devices as well as saving energy and bandwidth by computation offloading. MAMoC is designed to allow mobile devices to discover other surrounding devices over standard Local Area Network in infrastructural Wi-Fi using an access point, peer-to-peer Wi-Fi, and Bluetooth personal area networks as shown in Figure 1. The main goals of the framework are to allow mobile application developers to achieve a transparent automated offloading to multiple destination clouds (mobile clouds and public clouds) and device dynamic changes over the lifecycle of execution of an application.

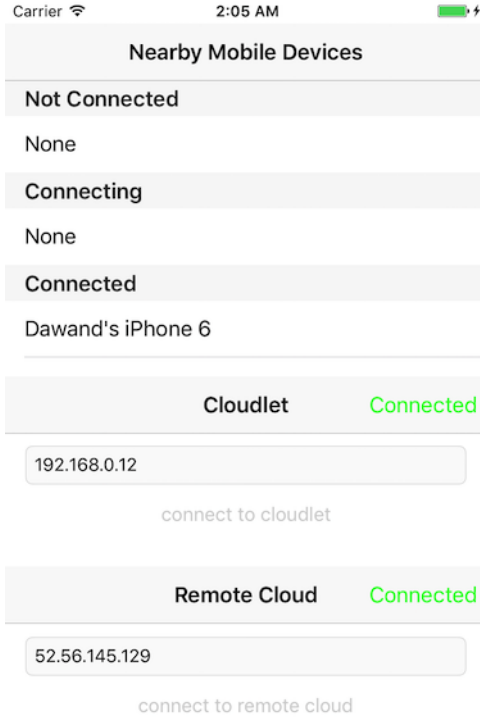


Figure 2. Service Discovery

A. Components

1) *Service Discovery*: After the framework is initialized, service discovery is performed. Each mobile device can advertise services and discover what services other nearby devices on the local network are offering. A browser object in a host device searches for peers which have an advertiser object. This can be done using infrastructural Wi-Fi, where the devices are connected to the same Access Point, Peer-to-peer Wi-Fi, or Bluetooth wireless technologies. An offloader device can also scan the local network for cloudlet servers and open a TCP socket communication channel for offloading requests. A standard user interface is developed to be used by the applications for service discovery and managing device connections as shown in Figure 2.

The mobile ad hoc cloud communication technique over LAN is based on Zero Configuration Network [18]. MultiPeer Connectivity (MPC) library [19] uses Zero Configuration Network technology to allow application developers to form short range sessions between nearby devices. The devices do not need a preconfigured network to exchange data with each other. There is no configuration needed because they can discover each other via multicast DNS (mDNS) [20]. mDNS is a service that resolves host names on a local network without the use of a central domain name server. After the connection is established between two devices, data can be exchanged between them in either reliable mode over TCP or unreliable mode over UDP. The reliable mode guarantees the delivery and avoids out of order packages so we have used this mode throughout our framework data transfers.

Table I
NOTATIONS

Symbol	Description
B_{man}	Benchmark workload result of mandelbrot set
B_{ftt}	Benchmark workload result of FFT
B	Average Benchmark workload result
B_{mob}	Benchmark score of a mobile device
B_{clet}	Benchmark score of a cloudlet
B_c	Benchmark score of a remote cloud
BL_{mob}	Battery level of a mobile device
M_{mob}	Available memory of a mobile device
M_{clet}	Available memory of a cloudlet
M_c	Available memory of a remote cloud
OS_{mob}	Offloading score of a mobile device
OS_{clet}	Offloading score of a cloudlet
OS_c	Offloading score of a remote cloud
P_{mob}	Computation speed of a mobile device
P_{clet}	Computation speed of a cloudlet
P_c	Computation speed of a remote cloud
RTT_{mob}	Network overhead and data transfer to mobile device
RTT_{clet}	Network overhead and data transfer to cloudlet
RTT_c	Network overhead and data transfer to cloud
TOS	Total offloading score of all the connected nodes

2) *Device Profiler*: This component collects real time information of the devices including hardware, software, and networking related context information. This allows the framework to have sufficient information about the connected devices. This eases the process of offloading decision making. Once a new device joins the framework, it goes through two profiling procedures. First, a collection of contextual information of the device is retrieved by the framework including hardware information such as number of CPU cores and processing speed, the type of the network (Wi-Fi or cellular), and battery level and state (whether it is charging or not). Second, a workload is sent over to the device after a successful connection establishment. The makespan result of running the workload is sent back to the framework. The collected information in both phases is then disseminated to all the connected devices in the framework to be used by the decision engine in the offloading process.

3) *Offloading Decision Engine*: The engine is fed with the metrics from device profiler component and it is used by the framework to check whether to offload a task to other devices or execute it locally. This component is only used in the partial offloading execution modes. Unlike full offloading, the framework needs to decide the candidate destinations for offloading the task, partition the task into a number of subtasks, and distribute them.

B. Offloading Models and Algorithms

When we model mobile computation offloading problems, we need to consider the time needed to execute computation locally, time needed to execute computation remotely (excluding the cost of data transfer), and time needed to move input data and results back and forth between the local and remote computers. We also need to collect the device context information in our preprocessing phase to model the offloading decision making process including the CPU power, memory, battery state and level of the connected devices.

The benchmark score will be calculated after the workloads are sent to the devices and the execution results are received. The quicker the CPU completes the tests, the higher the benchmark score. The workloads measure the instruction performance of the device by performing processor-intensive tasks that make heavy use of integer instructions. Initially, we create two types of workloads: compute-bound and memory-bound. Mandelbrot set [7] of an 800x800 pixels is used for the first type. The Fast Fourier Transform (FFT) [21] is used as a memory heavy workload. The workloads are executed 5 times and the average runtime score in GFlops is returned to MAMoC. Finally, the Benchmarking score of the device is calculated using equation 1:

$$B = (B_{man} + B_{fft})/2 \quad (1)$$

The computation power and memory of connected nearby mobile devices are initially collected. If the mobile device is not currently charging, then the offloading score is deducted by the amount of the used battery level of the device. The offloading score is then calculated using the following equation:

$$OS_{mob} = (B_{mob} + P_{mob} + M_{mob}) - RTT_{mob} - (100 - BL_{mob}) \quad (2)$$

Similar to mobile ad hoc cloud model, cloudlet and remote cloud modelling is a summation of their respective benchmark score, computation speed, and available memory subtracted by the data transfer cost.

$$OS_{clet} = (B_{clet} + P_{clet} + M_{clet}) - RTT_{clet} \quad (3)$$

$$OS_c = (B_c + P_c + M_c) - RTT_c \quad (4)$$

Given the offloading scores for all the offloadee device candidates, we calculate the total offloading score for any particular task execution.

$$TOS = |OS_{mob}| + |OS_{clet}| + |OS_c| \quad (5)$$

MAMoC collects the offloading scores of the local device running the mobile application and all the connected service providers. Algorithm 1 shows the process of collecting individual offloading scores calculated and received earlier to generate a dictionary of nodes and their corresponding offloading scores.

Algorithm 1 Offloading Score Algorithm

Input: connectedNodes

Output: nodeScores

```

1: nodeScores = [:]      ▷ A dictionary of nodes and their
   respective offloading scores
2: selfNode = getSelfNode()
3: localScore = getScore()      ▷ Based on equation 2
4: nodeScores.add(selfNode, localScore)
5: if connectedNodes > 1 then
6:   for node in connectedNodes do
7:     score = node.getScore() ▷ Based on equations 2-4
8:     nodeScores.add(node, score)
9:   end for
10: end if
11: return nodeScores

```

The node scores will then be sent to task partitioning algorithm to calculate the final task partitioning percentage for any given task.

Algorithm 2 Task Partitioning Algorithm

Input: nodeScores, totalScore

Output: partitioningResult

```

1: partitioningResult = [:] ▷ A dictionary of nodes and task
   allocation percentages
2: for (node, score) in nodeScores do
3:   partition = (score / totalScore) * 100
4:   partitioningResult.add(node, partition)
5: end for
6: return partitioningResult

```

C. Development Phase

We have used Swift¹ to implement both the client (iOS) and server (Linux) components of MAMoC. Swift is open sourced by Apple on December 2016. This has enabled developers to write Swift applications on broader range of platforms. To the best of our knowledge, this is the first complete Swift and iOS based framework in the mobile cloud computing research literature.

Containers are utilized to host the server components of MAMoC. Linux Containers (LXC) is a virtualization method for running multiple isolated Linux systems on a single machine. Docker² extends LXC to automate the deployment of applications inside software containers. We have used an already developed Swift Docker image to implement a server-side swift application to accept incoming requests from the mobile devices. The developed container provides an environment which is ready to be customised for other mobile applications. It provides a feature-rich yet lightweight execution environment for offloaded tasks.

We have run a couple of compute-intensive workloads to experiment the speed of computation on Docker containers as well as on native platforms. As the results show in 3, the

¹<https://swift.org>

²<https://www.docker.com/>

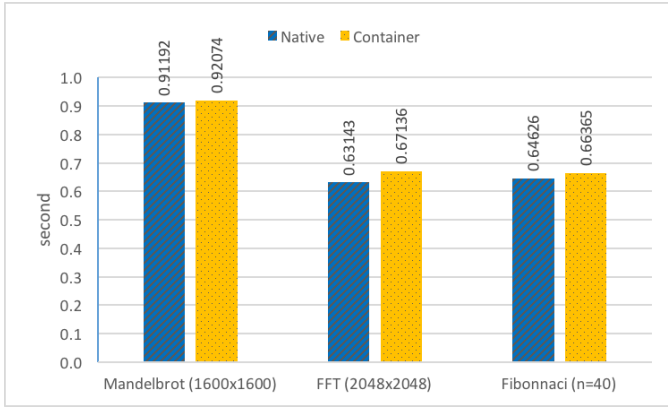


Figure 3. Container vs. native benchmarking

computation speed of containers are close to native platform performance within 1% region [22] [23]. To ease the creation of the container, we provide a Dockerfile that can automate the process of creating a container for the server-side part of MAMoC. The container is preconfigured with the necessary build environment for handling the client requests.

MAMoC allows other mobile application developers to adapt it for their own compute-intensive mobile applications. A developer must follow the following steps to configure an application to use our framework:

- 1) Add MAMoC to the mobile application: This process can be done manually by adding the files to the application directory or automatically using CocoaPods³ which is a dependency manager for Swift and Objective-C projects.
- 2) Define the task and its parameters: In this step, the developer has to override *initTask* method to provide any user input or additional data to be processed by the framework. In our text search example, the parameters are the text file and the search keyword which is entered by the user. In a face recognition application, the parameter would be an image of the face and so on.
- 3) Define the offloading sites: Even though MAMoC is designed to offload tasks to multiple sites by default, the developer can define where to execute the task. There are four choices to choose from: nearby mobile devices, cloudlet servers (local immobile), remote cloud servers, or automatic. The automatic option depends on the outcome of the models in the device profiler component and how the tasks are partitioned to target offloaders by the offloading decision engine.
- 4) Execute the job: the subtasks will be distributed accordingly to be processed in parallel on the target offloaders. The results are returned and merged together in the offloader device and presented to the user.
- 5) Set the timeout for the task reprocessing (optional): For the tasks which are not returned successfully by the resource providers, the task will be reprocessed locally

³<https://cocoapods.org/>

Table II
TESTBED EXPERIMENTATION

Node	Benchmark (in GFlops)	CPU (in GHz)	RAM (in GB)
Mobile (small)	1.09	1.3 (Dual)	1
Mobile (medium)	1.24	1.4 (Dual)	1
Cloudlet	2.56	2.5 (Quad)	16
Cloud (small)	2.32	2.4 (Single)	1
Cloud (medium)	2.94	2.8 (Quad)	7.5
Cloud (large)	3.02	2.8 (Octa)	15

in the device once the timeout duration is reached. The default value for this is 10 seconds. This adds fault tolerance feature to the applications developed on top of MAMoC.

The source code of MAMoC and a short documentation for setting up the different components in the framework is publicly available online at <https://github.com/dawand/MAMoC>

IV. EXPERIMENTATION

We have developed a mobile application to test the performance of the framework and showcase the different execution scenarios. We measure total completion time for the application in different settings. The application is executed in four different modes: local execution, nearby mobile devices, cloudlet, and remote cloud (with three different servers). We use four offloading scenarios: a full offloading and three types of partial offloading (workload sharing in a parallel manner) with different configurations. The application contains a Knuth-Morris-Pratt searching algorithm [24] to be performed on three different size text files. The large text file consists of 1,095,649 words, the medium text file contains 316,323 words, and the small text file contains 39,799 words. The files are stored in the mobile application running on the offloader device. The files need to be distributed through a wireless medium to the offloaders before the keyword search is conducted in the destination. Each execution was repeated ten times, such that averages could be calculated for more accurate results.

Our experimental testbed consists of two mobile devices (an iPhone 5 and an iPhone 6), one cloudlet, and three remote cloud instance types. The hardware specifications are shown in Table II. For the remote cloud instances, we have used three different Amazon Web Services instance types: small (t2.micro), medium (c4.xLarge), and large (c3.2xlarge). The two phones will be running the mobile application while cloudlet and remote cloud instances will be running the container described earlier.

In all our experiments, mobile (small) is the offloader and the rest of the nodes are service providers (offloaders). The offloading scenarios are:

- Full offloading: In full offloading mode, the execution is performed in the offloader and the final result is returned to the offloader. On the other hand, partial offloading mode only sends a part of the execution over to the offloader and performs the rest of the execution itself

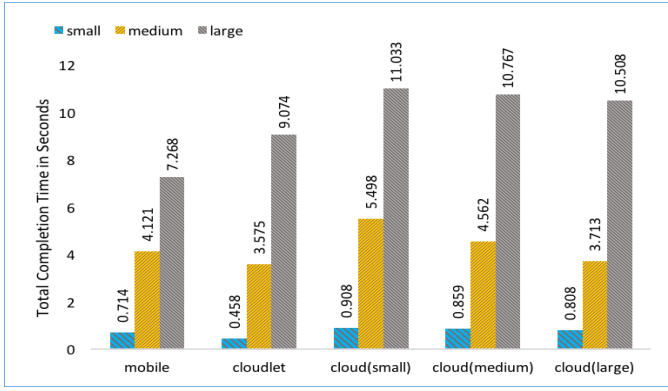


Figure 4. Full Offloading - The whole computation is offloaded to the offloadee

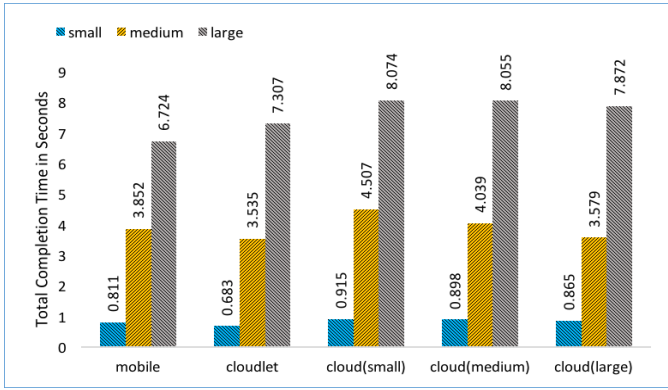


Figure 5. Partial Offloading (equal task distribution) - Local mobile device executes 50% of the task while the remaining 50% is offloaded

as we will observe in the other two scenarios. After the results are received, they are merged and stored in the offloader.

Our framework is specifically designed to support partial offloading. Nonetheless, we wanted to observe the completion times of full offloading executions and compare them to the decision engine results. We performed a full offloading experimentation on the offloaders separately using the three text files. Figure 4 shows the total completion time of running the application with different text file sizes. As we will see later, most of the completion time is the communication overhead that occurs during the transfer of the necessary data (the text file content) from the mobile device to the offloaders.

- Partial offloading: we perform our first set of partial offloading experiments with no help from MAMoC. The tasks are equally distributed among the connected nodes. In other words, if there is only one available offloadee, the workload is divided into two equivalent halves and distributed to them. Both devices then execute the workload in a parallel fashion. The local result and the result returned from the offloadee are then merged and stored locally. The results of running the same set of workloads as previous experiment are displayed in Figure 5.

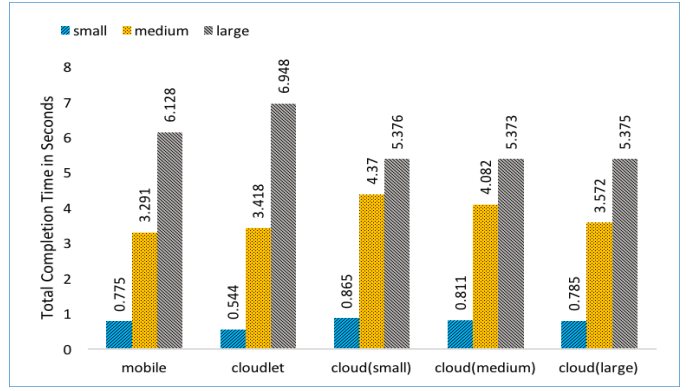


Figure 6. Partial Offloading with MAMoC (single-site) - The decision engine partitions the task for the local mobile device and a single offloadee

- Partial offloading with decision engine enabled: the decision engine uses the models presented earlier to calculate the percentage of the task which should be offloaded to any particular offloadee. The offloading scores of the offloader and the offloaders are investigated by the decision engine for the task partitioning process. In the case of offloading score being less than zero, no computation is offloaded to that offloadee. We wanted to observe the performance of our decision engine in both single-site and multi-site offloading scenarios. Figure 6 shows the single-site offloading scenario where the task is meant to be executed locally and a single offloadee. The task partitioning percentage and completion times of partitioned tasks for multi-site offloading scenario are displayed in Figure 7. It is worth noting that we have only used the large instance type of the remote cloud along with a cloudlet and a nearby mobile device as offloaders in multi-site offloading scenario. The complete set of results of all the experiments are shown in Table III.

V. RESULTS AND DISCUSSION

It is already shown in the literature that offloading does not always benefit the lower end devices [25]. In our first execution scenario, we send the text file to the destination and perform the search operation in the corresponding computation device. Despite few millisecond performance gains in the case of small text file, local execution is preferred to full offloading for medium and large text files. The partial offloading with equal partitioning of tasks among the local and external devices perform better in terms of reducing the overall network overhead occurrence in the previous execution scenario.

The single-site partial offloading produces better results than equal task distribution scenario in all of the offloading modes. It even produces a lower overall completion time for the cloudlet than the multi-site offloading scenario. This is due to the limited number of nearby mobile devices and cloudlets that are used in our testbed. Even though cloudlet-only full offloading performed better in both small and medium text

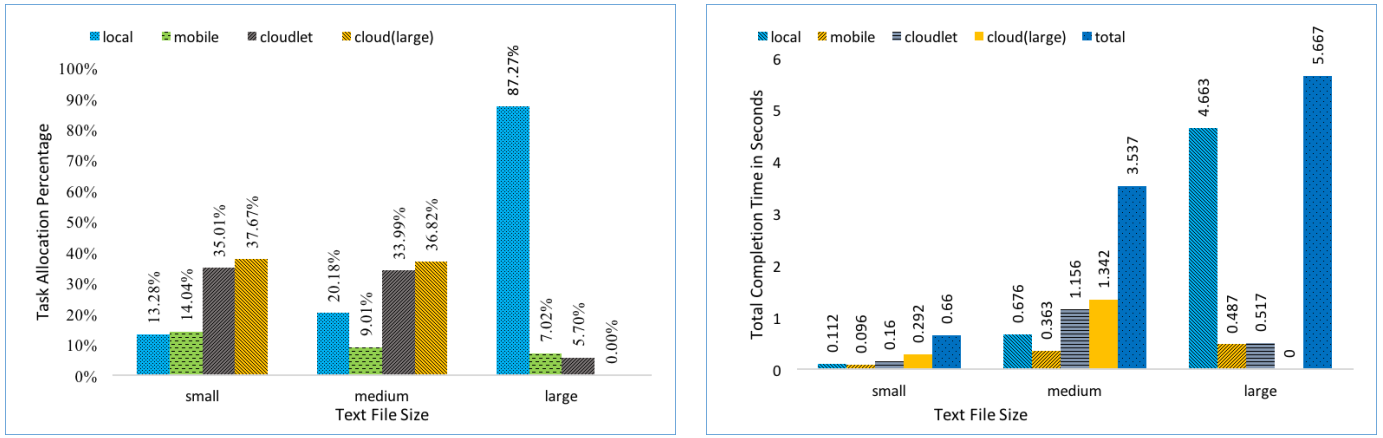


Figure 7. Multisite Partial Offloading - The decision engine partitions the task for the local mobile device and multiple offloaders

Table III
EXPERIMENTATION RESULTS

Text file size	Offloadee	Full offloading	Partial offloading (equal task partitioning)	Partial offloading (single)		Partial offloading (multi)		
		Completion time	Completion time	Task partitioning percentage	Completion time	Task partitioning percentage	Completion time	
Small	Mobile	0.714	0.811	48.60%	51.40%	13.28%	14.04%	0.659
	Cloudlet	0.458	0.683	27.90%	72.10%		35.01%	
	Cloud (small)	0.908	0.915	35.17%	64.83%		-	
	Cloud (medium)	0.859	0.898	28.57%	71.43%		-	
	Cloud (large)	0.808	0.865	26.45%	73.55%		37.67%	
Medium	Mobile	4.121	3.852	69.12%	30.88%	20.18%	9.01%	3.537
	Cloudlet	3.575	3.535	37.35%	62.65%		33.99%	
	Cloud (small)	5.498	4.507	55.25%	44.75%		-	
	Cloud (medium)	4.562	4.039	40.25%	59.75%		-	
	Cloud (large)	3.713	3.579	36.54%	63.46%		36.82%	
Large	Mobile	7.268	6.724	92.55%	7.45%	87.27%	7.02%	5.667
	Cloudlet	9.074	7.307	93.85%	6.15%		5.70%	
	Cloud (small)	11.033	8.074	100%	0%		-	
	Cloud (medium)	10.767	8.055	100%	0%		-	
	Cloud (large)	10.508	7.872	100%	0%		0%	

file size execution scenarios, large text file size offloading was more efficient in MAMoC (9.074 vs. 5.667 seconds) as a larger portion of execution is performed locally. This is mainly due to the delay caused by data transfer in full offloading execution. For the large text file size offloading scenarios, the network overhead occupies a much larger portion of

the total completion time than the time taken for the task execution. This shows a clear advantage of local execution over offloading when the needed computation is not as much. In short, the more computation and less data transfer is needed, offloading has more advantage.

Since we use Wi-Fi in the evaluation, the time of sending

files and receiving results has an impact on the overall performance, but if cellular networks or Bluetooth are used, the offloading time will surely increase.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we presented an adaptive multisite offloading framework that takes into consideration dynamic context changes in MCC environment and offloads computation to multiple offloaders including nearby mobile devices, cloudlets, and remote cloud servers. We evaluated the proposed framework, and results showed that it can provide suitable offloading decisions based on the current context of the local device and the external platforms. We developed a text search application and conducted experiments with three different text file sizes on different offloading scenarios. Our results present different insights into the factors that affect the offloading decision. Our future work includes increasing context parameters and the number of nearby devices and cloudlets in our experimental testbed. Our future experiments will not be confined to Wi-Fi and would explore other protocols such as Bluetooth Low Energy for nearby device-to-device communications and 4G for the mobile device and remote cloud communications. Moreover, we have built our framework with the assumption that the mobile user (offloader) stays in the same zone within the computation offloading duration. This cannot be guaranteed if the user moves away before the result from the destination is returned. The experiments in this paper were performed in a controlled setting. Enabling user mobility by live migrating the host mobile cloud containers is another future endeavor. An active research challenge is how to implement live migration in mobile cloud systems.

VII. REFERENCES

- [1] U. Drolia, R. Martins, J. Tan, A. Chheda, M. Sanghavi, R. Gandhi, and P. Narasimhan, "The case for mobile edge-clouds," in *2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing*, Dec 2013, pp. 209–215.
- [2] I. Yaqoob, E. Ahmed, A. Gani, S. Mokhtar, M. Imran, and S. Guizani, "Mobile ad hoc cloud: A survey," *Wireless Communications and Mobile Computing*, vol. 16, no. 16, pp. 2572–2589, 2016, wCM-16-0169.R1. [Online]. Available: <http://dx.doi.org/10.1002/wcm.2709>
- [3] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol. 8, no. 4, pp. 14–23, Oct 2009.
- [4] T. Verbelen, P. Simoens, F. De Turck, and B. Dhoedt, "Cloudlets: Bringing the cloud to the mobile user," in *Proceedings of the Third ACM Workshop on Mobile Cloud Computing and Services*, ser. MCS 12. New York, NY, USA: ACM, 2012, pp. 29–36. [Online]. Available: <http://doi.acm.org/10.1145/2307849.2307858>
- [5] C. Shi, V. Lakafosis, M. H. Ammar, and E. W. Zegura, "Serendipity: Enabling remote computing among intermittently connected mobile devices," in *Proceedings of the Thirteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '12. New York, NY, USA: ACM, 2012, pp. 145–154. [Online]. Available: <http://doi.acm.org/10.1145/2248371.2248394>
- [6] A. Mtibaa, K. A. Harras, and A. Fahim, "Towards computational offloading in mobile device clouds," in *2013 IEEE 5th International Conference on Cloud Computing Technology and Science*, vol. 1, Dec 2013, pp. 331–338.
- [7] N. Fernando, S. W. Loke, and W. Rahayu, "Computing with nearby mobile devices: a work sharing algorithm for mobile edge-clouds," *IEEE Transactions on Cloud Computing*, 2016.
- [8] Z. Sanaei, S. Abolfazli, A. Gani, and R. Buyya, "Heterogeneity in mobile cloud computing: taxonomy and open challenges," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 369–392, 2014.
- [9] K. Kumar, J. Liu, Y.-H. Lu, and B. Bhargava, "A survey of computation offloading for mobile systems," *Mob. Netw. Appl.*, vol. 18, no. 1, pp. 129–140, Feb. 2013. [Online]. Available: <http://dx.doi.org/10.1007/s11036-012-0368-0>
- [10] X. Zhang, A. Kunjithapatham, S. Jeong, and S. Gibbs, "Towards an elastic application model for augmenting the computing capabilities of mobile devices with cloud computing," *Mobile Networks and Applications*, vol. 16, no. 3, pp. 270–284, Jun 2011. [Online]. Available: <https://doi.org/10.1007/s11036-011-0305-7>
- [11] I. Giurgiu, O. Riva, D. Juric, I. Krivulev, and G. Alonso, "Calling the cloud: Enabling mobile phones as interfaces to cloud applications," in *Proceedings of the ACM/IFIP/USENIX 10th International Conference on Middleware*, ser. Middleware '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 83–102. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1813355.1813362>
- [12] B.-G. Chun, S. Ihm, P. Maniatis, M. Naik, and A. Patti, "Clonecloud: Elastic execution between mobile device and cloud," in *Proceedings of the Sixth Conference on Computer Systems*, ser. EuroSys '11. New York, NY, USA: ACM, 2011, pp. 301–314. [Online]. Available: <http://doi.acm.org/10.1145/1966445.1966473>
- [13] G. Huerta-Canepa and D. Lee, "A virtual cloud computing provider for mobile devices," in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*, ser. MCS '10. New York, NY, USA: ACM, 2010, pp. 6:1–6:5. [Online]. Available: <http://doi.acm.org/10.1145/1810931.1810937>
- [14] R. Kemp, N. Palmer, T. Kielmann, and H. E. Bal, "Cuckoo: A computation offloading framework for smartphones," in *MobiCASE*. Springer, 2010, pp. 59–79.
- [15] E. Cuervo, A. Balasubramanian, D.-k. Cho, A. Wolman, S. Saroiu, R. Chandra, and P. Bahl, "Maui: Making smartphones last longer with code offload," in *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '10. New York, NY, USA: ACM, 2010, pp. 49–62. [Online]. Available: <http://doi.acm.org/10.1145/1814433.1814441>
- [16] K. Sinha and M. Kulkarni, "Techniques for fine-grained, multi-site computation offloading," *2011 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 184–194, 2011.
- [17] D. Sulaiman and A. Barker, "Task offloading engine for heterogeneous mobile clouds," in *Proceedings of the 8th EAI International Conference on Mobile Computing, Applications and Services*. ACM, 2016.
- [18] S. Cheshire and D. Steinberg, *Zero Configuration Networking: The Definitive Guide*, ser. Definitive Guide Series. O'Reilly Media, 2006. [Online]. Available: <https://books.google.co.uk/books?id=R3jxPwQhUC>
- [19] A. D. Documentation. MultipeerConnectivity. [Online]. Available: <https://developer.apple.com/documentation/multipeerconnectivity>
- [20] S. Cheshire and M. Krochmal, "Multicast dns," Tech. Rep., 2013. [Online]. Available: <https://tools.ietf.org/html/rfc6762>
- [21] K. Katoh, K. Misawa, K.-i. Kuma, and T. Miyata, "Mafft: a novel method for rapid multiple sequence alignment based on fast fourier transform," *Nucleic acids research*, vol. 30, no. 14, pp. 3059–3066, 2002.
- [22] W. Felter, A. Ferreira, R. Rajamony, and J. Rubio, "An updated performance comparison of virtual machines and linux containers," in *2015 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, March 2015, pp. 171–172.
- [23] B. Varghese, L. T. Subba, L. Thai, and A. Barker, "Container-based cloud virtual machine benchmarking," in *2016 IEEE International Conference on Cloud Engineering (IC2E)*, April 2016, pp. 192–201.
- [24] D. E. Knuth, J. H. Morris, Jr, and V. R. Pratt, "Fast pattern matching in strings," *SIAM journal on computing*, vol. 6, no. 2, pp. 323–350, 1977.
- [25] K. Kumar and Y. H. Lu, "Cloud computing for mobile users: Can offloading computation save energy?" *Computer*, vol. 43, no. 4, pp. 51–56, April 2010.